Formalization of Learning Patterns Through SNKA

Mr Rajesh D

PhD Research Scholar
Department of Computer Science
Bharathiar University, Coimbatore, India

Dr. K.David

Assistant Professor in Computer Science
HH Rajah's College
Pudukottai – Trichy Dist, India

Abstract—**The Learning patterns found among the learners community is steadily progressing towards the digitalized world. The learning patterns arise from acquiring and sharing knowledge. More impact is found on the usage of knowledge sharing tools such as facebook, linkedin, weblogs, etc that are dominating the traditional means of learning. Since the knowledge patterns acquired through web unstructured data is insecure, it leads to poor decision making or decision making without a root cause. These acquired patterns are also shared to others which indirectly affect the trust patterns between users. In this paper, In order to streamline the knowledge acquisition patterns and their sharing means a new framework is defined as Social Networking based Knowledge Acquisition (SNKA) to formalize the observed data and the Dynamic Itemset Count (DIC) algorithm is tried for predicting the users about the usage of web content before and after the knowledge is acquired. Finally the rough idea in building a tool is also suggested.**

Keywords—Data Acquiring methods; Learning Patterns; Knowledge Management; Data Mining Tools

I. INTRODUCTION

The student's role in utilizing the web enabled systems has dominated their learning patterns in a day today life. Its growth rate is inevitable. In the process of gaining an updated knowledge in a faster mode, the social networking platform is used. This platform provides an easy structure to share information to anybody, at anyplace and at anytime. Even though, there are various educational tools and Environments, students prefer to learn through Digital web pages, web spaces and blogs.[7] The tremendous growth of information in blogs and other unauthorized web spaces leads to acquiring wrong information or unbiased information. These unstructured data becomes a real challenge for all data analyzers. the framework for acquiring specific valid data from various sources becomes the most important problem for all learners.

Most of the generic online users look into the social networking for finding friends of like-minded or domain-oriented or casual-time pass nature; to share their ideas, experiences and knowledge to others in turn they start learning or acquire knowledge indirectly. Still there is an issue of absoluteness and correctness in the acquired data.

This paper discusses about the social networking phases and their utilization among online learners. The facts behind the sharing of one's own ideas or forwarding the others post should be verified for their originality and trueness of the post. The Second part depicts the knowledge management frame work which is newly defined as Social Networking based Knowledge Acquisition (SNKA)[51] to formalize the observed data and the third part briefs about the supportive data mining algorithms which indirectly supports the last section describing the SNKA implementation through Dynamic Item set Counting algorithm.

II. SOCIAL NETWORKING BASED LEARNING PATTERNS

The Primary platform for the information and communication sharing is having a higher growth because the social networking tools. Hence, Social networking is defined as a virtual platform for people to interconnect and share among themselves. The users group of social networking is broadly categorized as known and unknown. The known group here depicts the known friends, friends of another friend and relatives in relationship whereas unknown group depicts the strangers who are either popularly known by all and people who we meet while walking through our life.[1][2] Social Networking is revolving around the world with different faces. Social Networking usage creates knowledge, shares information, explores data and finally depicts as a decision maker for many crucial factors. At a learner's juncture, this paper helps us to visualize the knowledge acquired through social networking and their analyzing efficiency towards effective decision making.

The reasons for making social networking more popular are: i) most attractive and faster communication source as it bridges a distance of less cost. Ii) All the posts quickly get response either through a Like or a comment. Iii) reduce tension at various levels of problems where one to one or one to many communication is regretted.

Apart from the above, the usage of social networking as a platform leads to gaining advantage with the factors like: i) Faster publishing ii) Quicker Impact and response, iii) Trust Building through originality and validity of the content posted among users, iv) Faster Publicity and Popularity and v) Quicker Reusability options.

A questionnaire was created as part of the study and circulated among various learners irrespective of age factor. The questionnaire was created as a online blog and also circulated manually. Table 1 describes the resultant of the questionnaire. The online filled questionnaire was overwhelming from the age group of 9 - 45 whereas none were there between 0 - 8 and 46 - 73. The traditional and manual filled forms had less impact and responses. Table2 depicts the impact of the students over their learning patterns dependent more on the digital web pages rather than the traditional learning patterns.

In order to prove the domination of social networking based learning six hypotheses were generated and studied.

They are:

i) To Analyze the users background information for understanding their browsing needs and abilities.
ii) To Check their sharing ethics and principals
iii) To Verify their priority in downloading patterns
iv) To Study the trust and behavior analysis
v) To Find the importance of designing a KM Framework for social networking web pages
vi) To Formalize the social networking web pages through SNKA and other data mining Algorithms

TABLE I.   ONLINE QUESTIONNAIRE RESPONSE

Questionnaire Analysis (Age Range)	
(Normative Study Method - Data Collection)	
Children upto 9 years of age	0
Adolescents 10-19 years	3
Adults 20-45 years	208
Middle age 46-60 years	21
Old > 60	2
Total	234

TABLE II.   IMPACT ON VARIOUS LEARNING PLATFORMS

Learning Pattern Observations	Impact
Social Networking as Learning Environment	66.5
General Browsing Based Learning	24.2
Traditional Learning Environment	9.3

The above impact of the study made to think about the decision making skills of the students who simply make their learning patterns as Social Networking based. A set of interview questions were forwarded to many Researchers, HR Managers and Administrators where knowledge is the base of Human Earnings. It was circulated to analyze their view about their employee's frequent usage and knowledge updating factors leading to decision making through social networking.

A few selected questions and their answers are represented in the table 1 and table2 respectively.

The following analysis shows that a serious cause over the learning patterns is indirectly supported by our environment. It has to be formalized for making a faster and efficient decision making efforts.

It also enables a strong bond between the sharing community irrespective of their groupings. The validity or originality behind the acquired and shared content makes a strong trust factor among users in the group which later gets multiplied within other users also.[5]

TABLE III.   Q1-SECONDARY DATA COLLECTION - INTERVIEW QUESTION

Interview Question : 1 Companies Expect Employees	Result
Frequent Update - Any means	35%
Allow SN usage	45%
Not Allow SN usage	20%

TABLE IV.   Q2-SECONDARY DATA COLLECTION - INTERVIEW QUESTION

Interview Question : 2 Usage of Social Networking	Result
Supports Decision Making Somewhere	55%
Does not Support Decision Making	25%
Not Allowed	20%

III. KNOWLEDGE MANAGEMENT

Knowledge is flourished everywhere in the form of information. The invention of knowledge from information is the process of knowledge management practices. Knowledge management is the process established to capture and use knowledge for the purpose of improving organization performance (Marakas, 1999) or individual's self interest. Knowledge management is emerging as the new discipline that provides the mechanisms for systematically managing the knowledge digitally. In order to have more promotions in the competitive world various knowledge acquisition tools are derived for faster knowledge generation and managing themselves. Knowledge management involves acquisition, generation, codification, storage, transfer, retrieval (Alavi and Leidner, 2001).

Knowledge is the asset of markets, products, technologies and organizations that a business owns or needs to own and which enable its business process to generate profits, and values. These knowledge management processes include:

i)   Developing knowledge,
ii) Preserving knowledge,
iii) Using knowledge, and
iv) Sharing knowledge.

The key elements of knowledge management are collaboration, content management and information sharing (Duffy, 2001). Collaboration refers to colleagues exchanging ideas and generating new knowledge. Common terms used to describe collaboration include knowledge creation, generation, production, development, use and organizational learning (Duffy, 2001). Content management refers to the management of an organization's internal and external knowledge using information skills and information technology tools. Terms associated with content management include information classification, codification, storage and access, organization and coordination (Alavi and Leidner, 2001

The incorporation of SNKA along with data mining algorithms will closely formalize the socio web pages where people can trust the posts published and could either be used for further purposes or as a decision making system. This SNKA was accomplished based on the generic knowledge acquisition cycle.

DEFINING SNKA CYCLE

The Social Networking based Knowledge Acquisition (SNKA)[51] Cycle depicts the process of absoluteness in knowledge building or knowledge gaining procedures. It provides a faster learning and structured acquisition patterns. The conversion process of unstructured data into structured data solves many deficiencies in learning patterns.

It's a chain of process where all social networking sites work under the same umbrella with a difference in typical human behavior. The Figure1 above depicts the SNKA Cycle which involves the process as:

Sent / Received: Information is sent or received through various sources of social networking tools.

Classify: Received information is to be classified for smoother usages.

Trueness: The validity of the received information is to tested through search engines or through trusted experts of the domain. The testing leads in finding the trueness such that a network of trustiness begins.

Re-usability: The verification process of the received information defines the status of reusability nature.

Store: Information is stored in the form of knowledge for futuristic usage after satisfying the verification and validation process

Share: After the successful completion of entire cycle, the acquired knowledge could be shared to others and hence this process surely leads to quality in decision making process at learner's view.
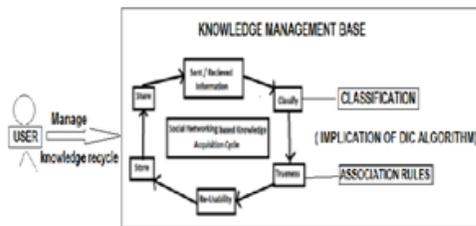

Fig. 1.   Implementation of SNKA with Data Mining Algorithms

IV. DATA MINING ALGORITHMS

The Primary data collected through the online mode was taken for analysis. Firstly the statistical tools were incorporated and in order to cross verify and finalize certain decisions pre - defined data mining algorithm tools like classification and clustering algorithms were tested in the obtained data.

Before selecting Dynamic Itemset Counting Algorithm, Apriori algorithm was analyzed. Since, DIC performed faster processing for samples data, We have undertaken taken to release the process without storing the counting of items sets that occurs frequently. [9][10]

The Generic DIC algorithm [Figure 2] was taken up as a case study that overrides as an alternative of Apriori algorithm in finding the association and frequency in web posts followed by their usage. the count in no. of revisit frequencies are used to determine the trustiness and the originality factors. Keeping these key factors as an input a rank indicator will be shown in future to users while and before they start sharing the posts they receive.The overall integration of classification and clustering relatively helps us in finding the correlation among the data sets. A FOAF was also taken for study (Friend Of A Friend).

ALGORITHM (Taken For Study)

1) Mark the empty itemset with a solid square. Mark all the 1-itemsets with dashed circles. Leave all other itemsets unmarked.

2) While any dashed itemsets remain:

a) Read M transactions (if we reach the end of the transaction file, continue from the beginning). For each transaction, increment the respective counters for the itemsets that appear in the transaction and are marked with dashes.

b) If a dashed circle's count exceeds minsupp, turn it into a dashed square. If any immediate superset of it has all of its subsets as solid or dashed squares, add a new counter for it and make it a dashed circle.

c) Once a dashed itemset has been counted through all the transactions, make it solid and stop counting it.

INTEGRATION OF DIC ALGORITHM WITH SNKA

Many Data mining methods were analyzed. The classification method was tried to classify the user profile classification based on their profile characterization. Their supportive algorithms are decision tree, Naïve Bayesian, Neural network and C&RT. The clustering methods were also mainly used for studying information retirval patterns. There were two ways to solve social networking data analysis. I) Graph Based requires study on nodes generated on the web structure and ii) test is based on clustering where new clusters were created based on the behavior of numerical data. Figure1 depicts the implication of SNKA with data mining algorithms. The user can manage his knowledge by following the SNKA framework to maintain his knowledge base stronger. This may impact more on their decisions making stronger.

Fig. 2.   Dynamic Itemset Counting (DIC)

[Figure3] depicts the process of creating a tool that rides on a browser especially on social networking blogs to provide a rank indicator for the posts that comes as inbox and takes care even after it goes out through the reference links

STEP1: USER LOGINS

STEP2: USER CHECKS FOR NEW UPDATES IN POSTS

STEP3: IT MAY EITHER BE NEW POSTS OR RETURN COMMENTS

STEP4: THE NEW MODEL CHECKS FOR THE RANKING

BASED ON: I) No. of Keywords/ Posts Searched, II) No. of Times Blog viewed, III) No. of Times further the Posted Content was downloaded or Referenced. These are calculated to identify the Frequent Item sets dynamically as the web posts are in the format of unstructured data.

STEP5: RANK INDICATORS PREDICTS THE USER AN INDICATION ABOUT THE POST THEY HAVE RECEIVED.

STEP 6: SHARABILITY OF THE POST IS FINALLY DEFINED.

Using this ranking the user may finally decide whether to forward the post blindly or not. This process also avoids the redundancy in data storage in web platforms.

It is also further may be linked with search engines for better ranking of web pages. Many other tools are provided by many developers like Graph Characterization Toolkit, Tweet hood, Meerkat, Netdriller, Hits/ISAC, D-Dupe and X-RIME are with certain limitations and different in implications.
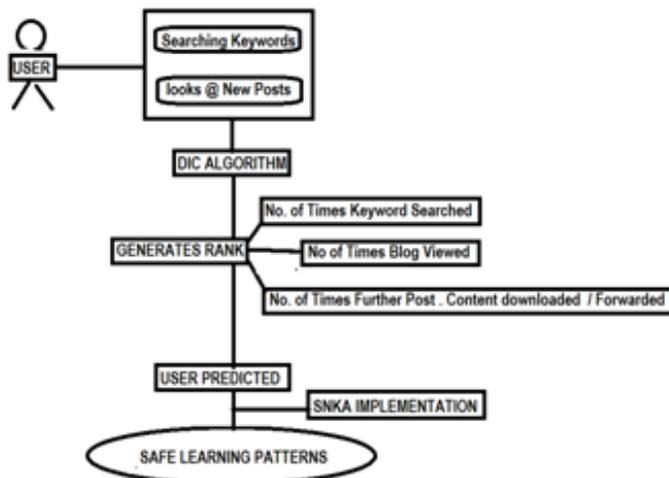

Fig. 3.   A Tool supporting SNKA

V. Limitations of the Study

This Study focused on the observations of the Indian digital learners alone. Only a few social networks are available for the study and hence, we have taken more focus on facebook, my blogs and linkedin. These social networking sites were mostly preferable among the users hence taken up for user analysis.

VI. Conclusion

The Study would help the digital learners to acquire valid knowledge patterns through SNKA and Data mining tools. This study helps to find the originality, trustfulness and the

sharing patterns of any content through knowledge management.. This paper helps to create a SNKA for the unstructured data. The sample datasets of unstructured social networking web pages were analyzed through empirical data collected by qualitative mode. As a further step, we would develop a tool to retrieve a fine tuned knowledge acquisition patterns as a trial and error method.

References

[1]   Adamic, Lada, Buyukkokten,Orkut, and Eytan Adar. (2003). A social network caught in the Web. *First Monday*, 8 (6).

[2]   Ahn, June. (2012). Teenagers' experiences with social network sites: Relationships to bridging and bonding social capital. *The Information Society*, 28(2), 99-109.

[3]   Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., and Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization.. *Psychological Science*, 21, 372-374.

[4]   Bo Fu, and Declan O\'Sullivan. (2007). Trust Management in Online Social Networks. *In Proceedings: the 7th IT&T Conference, - Digital Convergence in a Knowledge Society*. (pp. 3-12).

[5]   Bonneau, Joseph, Anderson, Jonathan, and Danezis, George. (2009). Prying Data Out of a Social Network. To be presented at the 2009 International Conference on Advances in Social Network Analysis and Mining, 2009.

[6]   Fergus Toolan "Web Mining" Intelligent Information Retrieval, Group University College, Dublin

[7]   Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 2008

[8]   Pan F, Cong G, Tung AKH, Yang J, Zaki M (2003) CARPENTER: finding closed patterns in long biological datasets. In: Proceeding of the 2003 ACM SIGKDD international conference on knowledge discovery and data mining (KDD'03), Washington, DC, pp 637–642

[9]   S. Brin, R. Motwani, et al, "Dynamic Itemset Counting and Implication Rules for Market Basket Data," SIGMOD Record (SCM Special Interset Group on Management of Data), 26,2, 1997.

[10]   Greenhow, Christine. (2011). Online social networks and learning. *On the Horizon*, 15 (1), 4-12. Learning, youth, informal and formal education

[11]   Griffiths, Marie, and Ben Light. (2008). Social networking and digital gaming media convergence: Classification and its consequences for appropriation. *Information Systems Frontiers*, 10 (4), 447 – 459

[12]   Wasko, M., and Faraj, S. (2005). "Why Should I Share? Examining Knowledge Contribution in Electronic Networks of Practice," MIS Quarterly 29 (1), pp 1-23.

[13]   Davenport, Thomas H. (1994). "Saving its Soul: Human Centered Information Management". *Harvard Business Review* **72** (2): 119–131.

[14]   Bordia, P., Irmer, B. E., & Abusah, D. (2006). Differences in sharing knowledge interpersonally and via databases: The role of evaluation apprehension and perceived benefits. European Journal of Work and Organizational Psychology, 15(3), 262−280

[15] Reagans, R., & McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. Administrative Science Quarterly, 48(2), 240−267

[16] Lin, C. -P. (2007a). To share or not to share: Modeling knowledge sharing using exchange ideology as a moderator. Personnel Review, 36(3), 457−475

[17] Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. MIS Quarterly, 25(1), 107−136

[18] Argote, L., McEvily, B., & Reagans, R. (2003). Managing knowledge in organizations: An integrative framework and review of emerging themes. Management Science, 49(4),571−582

[19] Jump up to: *a b c d e f g* "Introduction to Knowledge Management". University of North Carolina at Chapel Hill. Retrieved 18 April 2013.

[20] Sveiby, Karl. The facts about knowledge. *In* Knowledge management: Cultivating knowledge professionals, edited by Al-Hawamdeh, Suliman.Chandos Publishing, Oxford, 2003.pp. 32-33.

[21] Davenport, Thomas H. & Prusak, Larry. Knowledge management: Executive brief. *In* Knowledge management and management learning: Extending the horizons of knowledge-based management, edited by Baets, Walter, 2nd Ed., Springer, USA, 2005. pp. 69-70.

[22] Despres, Charles & Chauvel, Daniele. *In* Knowledge horizons: The present and the promise of knowledge management, 2nd Ed., Butterworth-Heinemann, Boston, 2001. pp. 106-07.

[23] Balmisse, Gilles, *et al*. Technology trends in knowledge management tools. *Int.. J. Know.Manage.*, 2007, **3**(2), 34-39.

[24] Firestone, Joseph M. & McElroy, Mark W. Key issues in the new knowledge management. Butterworth-Heinemann, Burlington, 2003.

[25] Miller, Ron. The evolution of knowledge management's *Content*, 2005, **28**(11), 22-28.

[26] Vizcaino, Aurora. Developing knowledge management systems from a knowledge based and multi agent approach. *Int. J. Know. Manage.,* 2007, **3**(4), 33-42

[27] Emilio ferrara, Pasquale de meo and giacomo fiumara, Robert Baumgartner, "Web data extraction, applications and Techniques: a survey", Acm computing surveys, vol. V, no. N, july 2012, pages 1{54

[28] Shian-Hua Lin, Jan-Ming Ho, "Discovering Informative Content Blocks from Web Documents" , *SIGKDD ''02*, July 23-26, 2002, Edmonton, Alberta, Canada.

[29] P.Sudhakar, G.Poonkuzhali, R.Kishore Kumar, *Member IAENG,"*Content Based Ranking for Search Engines", Proceedings of the International MultiConference of Engineers and Computer Scientists 2012 Vol I, IMECS 2012, March 14 - 16, 2012, Hong Kong

[30] Neha Gupta, Dr. Saba Hilal, " A Heuristic Approach for Web Content Extraction", International Journal of Computer Applications (0975 – 8887), Volume 15– No.5, February 2011

[31] G. Poonkuzhali, K.Thiagarajan, K.Sarukesi and G.V.Uma "Signed Approach for Mining Web Content Outliers" ,World Academy of Science, Engineering and Technology 56 ,,09

[32] P. Sivakumar, R. M. S Parvathi, "An Efficient Approach of Noise Removal from Web Page for Effectual Web Content Mining", European

[33] Journal of Scientific Research ISSN 1450-216X Vol.50 No.3 (2011), pp.340-351

[33] G. Poonkuzhali, K.Thiagarajan, K.Sarukesi and G.V.Uma "Elimination of Redundant Links in Web Pages" ,World Academy of Science, Engineering and Technology 52 2009

[34] Jaroslav Pokorny, Jozef Smizansky, "page content rank: an approach to the web content mining", Charles University, Faculty of Mathematics and Physics , Malostranské nám. 25, 118 00 Praha 1, Czech Republic

[35] Fergus Toolan "Web Mining" Intelligent Information Retrieval, Group University College, Dublin

[36] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 2008

[37] W. Bruce Croft, Donald Metzler, and Trevor Strohman, *Search Engines: Information Retrieval in Practice*, Addison Wesley, 2009

[38] Peter Brusilovsky, Jae-wook Ahn and Edie Rasmussen "Teaching Information Retrieval With Web-based Interactive Visualization" 25 July 2010

[39] Prahaladrao.M "Knowledge Management", Defence Electronics research laboratory Hyderabad.

[40] C. Romero, s Ventura," Educational Data mining: as urvey 1995-2005", Expert Systems with Applications 33 (2007) 135−146

[41] Steve Dale ,"Engaging the Social Web for Personal Knowledge Management (PKM)",March 5, 2012

[42] Almahdi Mohammed Ahmed., Nortia Md. Norwawi., Wan Hussain and Wan Ishak. (2009). Identifying Student and Organisation Matching Pattern Using Apriori Algoritham for Practicum Placement. International Conference on Electrical Engineering and Informatics. 5 – 7 August 2009, Selngor, Malaysia.

[43] Berry, M.J.A., and Linoff, G. (2000). Mastering Data Mining. New York: Wiley.

[44] Cherkassky,V.S., and Mullar,F. (1998). Learning from Data: Concepts, Theory and Methods. New York: Wiley.

[45] David Hand, Heikki Mannila and Padhraic Smyth. (2009). Principles of Data Mining. New Delhi: PHI Learning Private Limited.

[46] Grenander, U. (1996). Elements of Pattern Theory. Baltimore, MD: Jpohns Hopkins University Press.

[47] Kilduff,M., & Tsai, W. (2003). Social networks and Organisations. London/New Delhi: Sage Publications

[48] Sabherwal, R., & Sabherwal, S. (2005). Knowledge Management Using Information Technology: Determinants of Impact on Firm Value. Decision Science, 36(4), 531-567.

[49] Wasserman, S., & Faust, K. (1994). Social Network Analysis: Methods and Application. New York: Cambridge University Press.

[50] Wiig, K., (1995). Knowledge Management, Methods, Practical Approaches to Managing Knowledge. USA: Schema Press.

[51] Rajesh Doss, Dr.K.David, *"A Study and Review on generic comparison of Information Retrieval Methods*", International Journal on Advanced Computer Technology, Volume 2 Journal 3, June 2013.